Genome-wide association studies identify 77 loci for suicidality and provide novel biological insights

Sarah M.C. Colbert BA* [1,2,3,4], the Psychiatric Genomics Consortium Suicide Working Group, Douglas Ruderfer PhD [5,6,7], Anna R. Docherty PhD [8,9,10], Niamh Mullins PhD* [1,2,3,4]

[1] Department of Psychiatry, Icahn School of Medicine at Mount Sinai
[2] Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai
[3] Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai
[4] Department of Artificial Intelligence and Human Health, Icahn School of Medicine at Mount Sinai
[5] Department of Biomedical Informatics, Vanderbilt University Medical Center
[6] Vanderbilt Genetics Institute, Department of Medicine, Division of Genetic Medicine, Vanderbilt University Medical Center
[7] Department of Psychiatry and Behavioral Sciences, Vanderbilt University Medical Center
[8] Department of Psychiatry, University of Utah Spencer Fox Eccles School of Medicine
[9] Huntsman Mental Health Institute
[10] Clinical and Translational Science Institute & the Center for Genomic Medicine, University of Utah

*Correspondence: niamh.mullins@mssm.edu and sarah.colbert@icahn.mssm.edu

ABSTRACT

Suicidality is heritable and a leading cause of global morbidity and mortality, yet its biological etiology remains largely elusive. We conducted multi-ancestry genome-wide association study meta-analyses of suicidal ideation (259,747 cases), suicide attempt (64,993 cases), suicide death (9,197 cases), and suicidal behavior (suicide attempt/ death, 75,300 cases), across 54 cohorts (e.g., the Psychiatric Genomics Consortium, Million Veteran Program, UK Biobank). We identified 77 significant loci across meta-analyses, including 59 previously unreported for suicidality. SNP-based heritability ranged from 2.0-6.7% and there were strong, yet incomplete, genetic correlations between suicidality phenotypes (0.70-0.88). Fine-mapping prioritized 27 putative causal SNPs and 20 credible genes. Enrichment analyses implicated synaptic pathways and neuronal populations predominantly in subcortical brain regions (e.g., amygdala excitatory, medium spiny, hippocampal CA1-3). Together, these findings establish suicidality as a polygenic set of traits with both shared and distinct genetic influences, providing a foundation for future studies of suicide biology and etiology.

MAIN

Over 700,000 individuals die by suicide each year, making suicide a leading cause of death worldwide, and many more attempt suicide[1–4]. Suicidality encompasses a phenotypically heterogenous spectrum of thoughts, behaviors, and outcomes[5–10], ranging from suicidal ideation (SI) to suicide attempt (SA) and suicide death (SD). SI reflects thoughts of engaging in suicide-related behavior, SA involves non-fatal self-directed potentially injurious behavior with any intent to die, and SD represents death caused by self-injurious behavior with any intent to die[11].

Twin and family studies suggest that genetic factors influence suicidality, with heritability estimates ranging from 30-55%[12]. While suicidality often co-occurs with psychiatric disorders[13–15], the heritability of suicidality phenotypes remains significant even when accounting for comorbid psychiatric disorders, supporting the notion that while there is partial genetic overlap, suicidality also involves unique genetic influences not entirely explained by psychiatric disorders[16–18]. Genome-wide association studies (GWAS) have begun to uncover the polygenic architecture of suicidality phenotypes, previously identifying five loci for SI, 12 for SA (in some studies combining SA and SD), and two for SD[19–21]. However, their genetic architecture and underlying biology remain incompletely understood, largely due to modest sample sizes and insufficient ancestral diversity.

To address these gaps, the Psychiatric Genomics Consortium Suicide Working Group (PGC SUI) reports here the largest and most ancestrally diverse GWAS meta-analyses of SI (N_{Cases} = 259,747), SA (N_{Cases} = 64,993), and SD (N_{Cases} = 9,197) to date. In addition to these individual phenotypes, we also analyzed an aggregated phenotype, suicidal behavior (SB; N_{Cases} = 75,300), defined here to include SA or SD as per previous GWAS[17,20] to maximize sample size. Across these four suicidality phenotypes, we identify 77 genome-wide significant (GWS) loci, 59 of which are previously unreported, and further characterize the biological underpinnings of suicidality, providing insights to inform future prevention and intervention efforts.

RESULTS

GWAS meta-analyses reveal heritability and loci associated with suicidality

We conducted GWAS meta-analyses for SI, SA, SD, and SB across 54 studies worldwide. Cases and controls were defined using our published PGC SUI phenotyping protocol from Colbert et al.[5] which details the specific criteria and rationale and is briefly described in Table 1. Cohort-specific phenotype definitions are available in the Supplementary Note.

Table 1: Suicidality phenotype definitions and case and control criteria summarized from PGC SUI Phenotyping Protocol

Phenotype (Abbreviation), Definition	**Case criteria**	**Control criteria**	**Aggregated phenotype (Abbreviation)**	
Suicidal Ideation (SI), Thoughts of engaging in suicide-related behavior.	Evidence of lifetime SI specifically (SA does not automatically meet criteria). SI must be distinguished from thoughts of self-harm. Psychiatric disorder may/ may not be present.	No history of SI. Psychiatric disorder may be present. Exclusion criteria: history of SA or SD, psychiatric diagnosis AND missing SI information.	Suicidality	
Suicide Attempt (SA), A non-fatal self-directed potentially injurious behavior with any intent to die. May/ may not result in injury.	Evidence of lifetime SA specifically (SD does not automatically meet criteria). SA must be distinguished from non-suicidal self-injury. Psychiatric disorder may/ may not be present.	No history of SA. Psychiatric disorder or SI may be present. Exclusion criteria: SD, psychiatric diagnosis AND missing SA information.		Suicidal Behavior (SB)
Suicide Death (SD), Death caused by self-directed injurious behavior with any intent to die as a result of the behavior.	Evidence of SD. Psychiatric disorder may/ may not be present.	No SD, SA (since most controls are living) or deaths of undetermined intent. Psychiatric disorder or SI may be present. Exclusion criteria: psychiatric diagnosis AND missing SA information.		

Table 2: Sample sizes and numbers of cohorts for each suicidality phenotype by genetic ancestry or population.

	Suicidal ideation (SI)			Suicide attempt (SA)			Suicide death (SD)			Suicidal behavior (SB)*		
Genetic Ancestry or Population	N Cohorts	N Cases	N Controls	N Cohorts	N Cases	N Controls	N Cohorts	N Cases	N Controls	N Cohorts	N Cases	N Controls
African	7	52,615	188,733	8	6,710	128,960	-	-	-	8	6,753	128,960
South Asian	1	243	1,471	1	82	1,632	-	-	-	1	82	1,632
East Asian	5	8,026	30,999	5	2,169	23,519	1	746	14,049	6	2,918	37,568
European	20	176,147	1,010,300	29	53,919	1,063,988	5	7,584	652,070	30	62,533	1,090,754
Latin American	4	22,716	78,440	3	2,113	50,938	1	867	2,043	4	3,014	52,981
Multi-ancestry	37	259,747	1,309,943	46	64,993	1,269,037	7	9,197	668,162	49	75,300	1,311,895

*The SB phenotype enabled inclusion of cohorts where both SA and SD data were available but the number of SD cases alone was insufficient to contribute to the SD only GWAS meta-analysis. As a result, the total SB case count does not equal the simple sum of SA and SD cases. More details are provided in the Methods and Supplementary Table 1.

Inverse-variance-weighted fixed-effects meta-analyses were performed both within and across genetic ancestry or population groups (AFR = African, SAS = South Asian, EAS = East Asian, EUR = European, and LA = Latin American). Table 2 provides sample sizes for each meta-analysis, and Supplementary Table 1 details cohort-specific sample sizes.

All meta-analyses were assessed for inflation using linkage disequilibrium score regression (LDSC) with ancestry/population-matched reference panels[22]. LDSC intercepts ranged from 0.99-1.06, indicating little to no test statistic inflation (Supplementary Table 2). We observed significant liability scale SNP-heritability (h^2_{SNP}; $p < 0.0029$) in the EUR GWAS for SA (6.6%), SI (3.0%), SD (5.1%), and SB (6.7%) (Supplementary Table 2). h^2_{SNP} estimates in the AFR, EAS, and LA GWAS were similar, but showed larger standard errors due to limited sample sizes (Supplementary Table 2).

Across all GWAS meta-analyses (Supplementary Tables 3-10), we identified 77 unique genome-wide significant loci ($p < 5\times10^{-8}$) associated with at least one suicidality phenotype

(Figure 1, Supplementary Table 11). These included 18 loci previously reported for suicidality and 59 loci not overlapping with any locus previously reported as GWS for any suicidality phenotype[17–21,23,24].

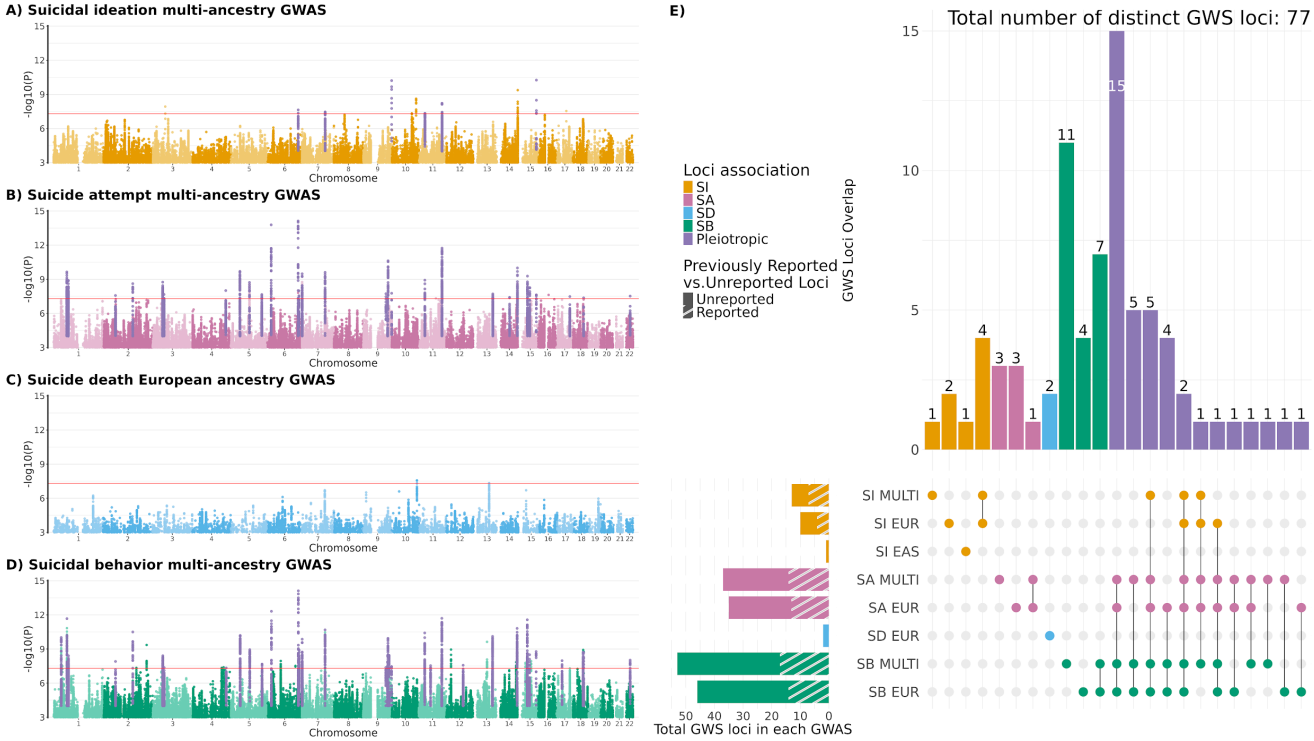

Figure 1. Manhattan and UpSet plots for GWAS meta-analyses of each suicidality phenotype. In panels A-D, the x-axis shows chromosomal position and the y-axis shows significance of the association as -log10(P). The red line indicates the genome-wide significance threshold ($p = 5 \times 10^{-8}$). Purple peaks indicate pleiotropic loci (i.e., loci significantly associated with more than one phenotype). **A)** Manhattan plot for the multi-ancestry GWAS of suicidal ideation (SI, yellow). **B)** Manhattan plot for the multi-ancestry (MULTI) GWAS of suicide attempt (SA, pink). **C)** Manhattan plot for the European (EUR) ancestry GWAS of suicide death (SD, blue). **D)** Manhattan plot for the multi-ancestry GWAS of suicidal behavior (SB, green). **E)** UpSet plot depicting the 77 distinct GWAS/phenotype-specific (SI = yellow, SA = pink, SD = blue, SB = green) and pleiotropic (purple) genome-wide significant (GWS) loci associated with the various suicidality GWAS. Vertical bars represent the number of loci associated with each combination of GWAS, as indicated in the intersection matrix below. For example, 11 loci were identified only in the SB multi-ancestry GWAS, while 15 loci overlapped between the SB multi-ancestry, SB EUR, SA multi-ancestry, and SA EUR GWAS. Horizontal bars represent the number of loci identified in the respective GWAS and full shading represents previously unreported loci, while diagonal shading represents previously reported loci. For example, the SD EUR GWAS identified two genome-wide significant loci, both of which have not been reported in previous suicidality GWAS. Full information for each locus and its overlap is provided in Supplementary Table 11. Results for genome-wide significant loci in each GWAS are reported in Supplementary Tables 3-10 and forest and area plots of the lead SNPs at GWS loci are provided in

Supplementary Figures 1-16. Ancestry/population-specific Manhattan plots are presented in Extended Data Figures 1-4.

The multi-ancestry SI GWAS identified 13 GWS loci (Supplementary Table 3, Figure 1A). The EUR SI GWAS detected seven of these and three additional loci (Supplementary Table 4, Extended Data Fig. 1C). The EAS SI GWAS identified one additional previously unreported locus (Supplementary Table 5, Extended Data Fig. 1B). No GWS loci were identified in the AFR or LA SI GWAS (Extended Data Fig. 1).

For SA, we identified 37 GWS loci in the multi-ancestry GWAS (Supplementary Table 6, Figure 1B) and 35 loci in the EUR GWAS (Supplementary Table 7, Extended Data Fig. 2C), with 27 overlapping between the two (Figure 1E). Of the 45 total loci associated with SA, nine overlapped with loci identified in the SI GWAS (Figure 1E, Supplementary Table 11). No GWS loci were detected in the AFR, EAS, or LA GWAS of SA (Extended Data Fig. 2).

For SD, no loci reached genome-wide significance in the multi-ancestry GWAS (Extended Data Fig. 3). However, the EUR GWAS identified two additional loci (Supplementary Table 8, Figure 1C) not previously associated with any published suicidality GWAS, and not identified in the current SI, SA, or SB GWAS. These loci included one in the *ARMS2* gene on chromosome 10 and one intergenic locus on chromosome 13.

In the SB GWAS, we detected 53 GWS loci in the multi-ancestry analysis (Supplementary Table 9, Figure 1D) and 46 loci in the EUR analysis (Supplementary Table 10, Extended Data Fig. 4), with 39 of these loci overlapping between the two analyses (Figure 1E, Supplementary Table 11). Notably, the SB GWAS identified 22 additional loci which did not reach genome-wide significance in the SA GWAS. Furthermore, nine of the SB loci overlapped with SI GWS loci. No GWS loci were detected in the AFR, EAS, or LA GWAS of SB (Extended Data Fig. 4).

Although no GWS loci were identified for SI, SA, and SB in AFR and LA GWAS, or for SA and SB in EAS GWAS, several lead SNPs from the multi-ancestry GWAS showed concordant effect directions in these groups (Supplementary Tables 12-14).

Shared and distinct genetic architectures of suicidality and related phenotypes

Unless otherwise specified, all post-GWAS analyses were based on multi-ancestry GWAS results to maximize statistical power. We estimated genome-wide genetic correlations (r_g) between the suicidality phenotypes using LDSC[25]. SA and SB showed near-complete genetic correlation (r_g = 1.02, se = 0.003, false discovery rate (FDR) p = 3.96×10^{-307}; Supplementary Table 15). Given their substantial sample overlap and resulting genetic similarity, we prioritized the SB GWAS over the SA GWAS for downstream analyses because it had greater power (larger sample size, higher mean χ^2, and more significant hits). Among SI, SD, and SB, the strongest genetic correlation was observed between SI and SB (r_g = 0.88), followed by lower correlations between SD and SB (r_g = 0.73) and between SD and SI (r_g = 0.70) (all FDR p < 0.05) (Figure 2, Supplementary Table 15). All three genetic correlations were significantly below 1 (all p < 8.01×10^{-5} for tests for r_g < 1), indicating partially distinct genetic architectures among these phenotypes. Each phenotype also displayed a moderate genetic correlation with non-suicidal self-injury (r_g = 0.39-0.56, all FDR p < 0.05), indicating a potentially unique genetic component specific to suicidal intent.

We also calculated the genetic correlation of SI, SD, and SB with relevant psychiatric and behavioral phenotypes (Supplementary Table 15). Among psychiatric disorders, suicidality phenotypes had the strongest genetic correlations with borderline personality disorder (r_g = 0.61-0.86) and internalizing disorders, including post-traumatic stress disorder (PTSD: r_g = 0.46-0.83), anxiety (r_g = 0.47-0.84), and major depressive disorder (MDD: r_g = 0.54-0.90) (all FDR p < 0.05) (Figure 2, Supplementary Table 15). Across most psychiatric disorders, and especially for internalizing disorders, genetic correlations with SI and SB were consistently higher than those with SD (Figure 2). Genetic correlations with substance use revealed further distinctions. All suicidality phenotypes showed generally modest genetic correlations with consumption-related substance use, whereas genetic correlations with problematic substance

use were significantly higher (all tests for significant differences FDR p<0.05; Supplementary Tables 15-16). SB exhibited significantly higher genetic correlations than SI with externalizing behavior (r_g = 0.57 vs. 0.41), problematic substance use phenotypes (e.g., addiction risk factor: r_g = 0.65 vs. 0.52), and attention-deficit/hyperactivity disorder (ADHD: r_g = 0.67 vs. 0.55) (all tests for significant differences FDR p<0.05; Supplementary Table 15, Supplementary Table 17), indicating that externalizing liability potentially contributes specifically to the etiology of and progression to SB. Interestingly, however, SI and SB did not differ in their genetic correlations with impulsivity (Figure 2, Supplementary Table 15).

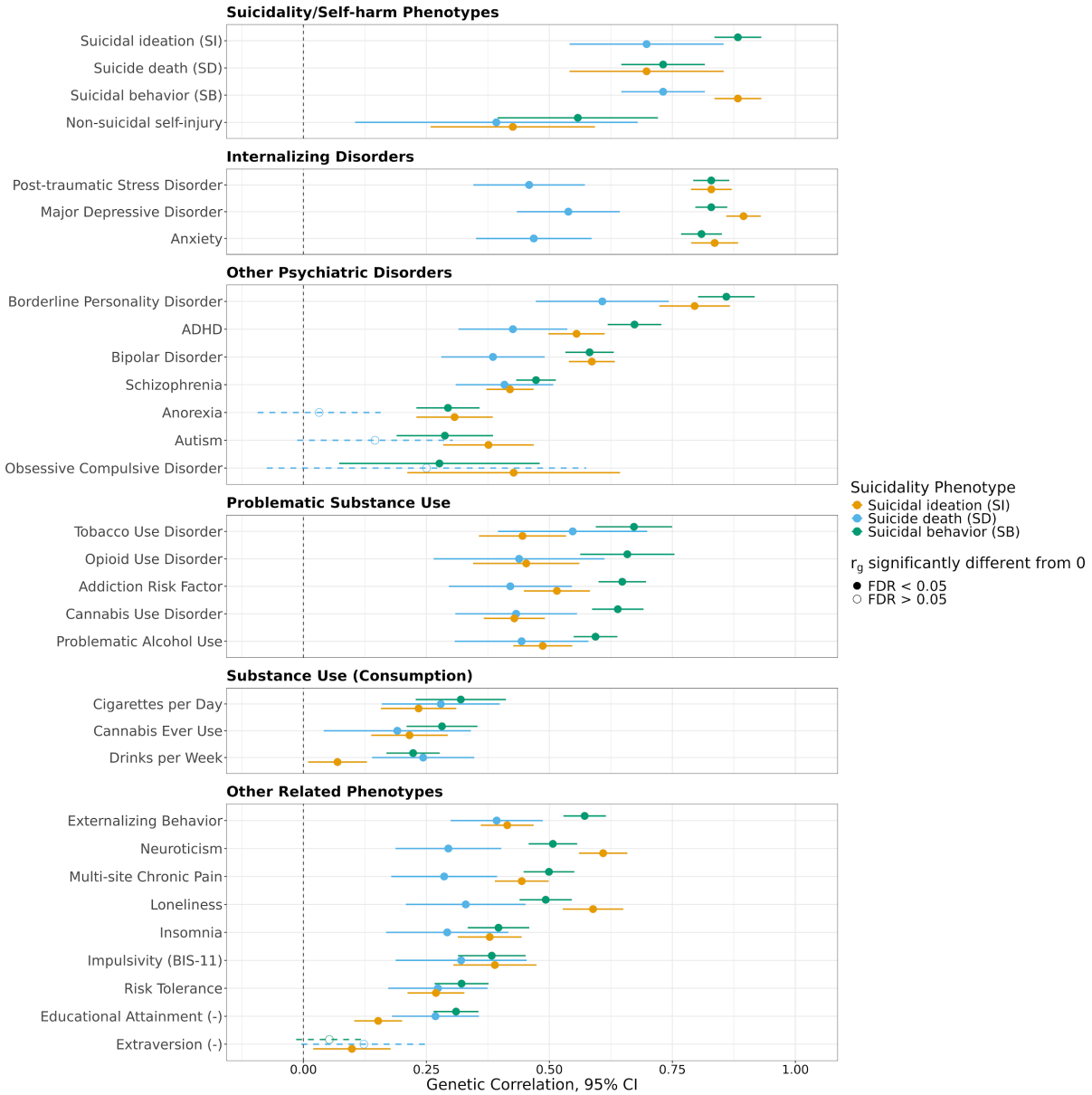

Figure 2: Genetic correlations (r_g) between suicidality, psychiatric, and behavioral

phenotypes. The x-axis shows genetic correlation estimates (points) with 95% confidence intervals (error bars). Points are colored according to suicidality phenotype (yellow = suicidal ideation, blue = suicide death, green = suicidal behavior). Shaded points indicate an r_g significantly different from zero (false discovery rate (FDR) < 5%). (-) indicates reverse-coded phenotypes. Additional related phenotypes tested but not shown on the figure are available in Supplementary Table 15.

Polygenic risk scores capture phenotypic variance in suicidality

Using PRS-CSx[26], we generated multi-ancestry polygenic risk scores (PRS) for suicidality phenotypes in 25 cohorts, each excluded in turn from the discovery GWAS. Within each cohort, we evaluated PRS performance using variance explained on the liability scale (R^2), odds ratios (ORs) comparing individuals in the top versus middle PRS quintiles, and area under the receiver operating characteristic curve (AUC) values (Supplementary Tables 18-21 and Supplementary Figures 17-19). Then, we calculated effective sample size (N_{eff})-weighted averages of these metrics within and across the AFR, EAS, EUR, and LA groups.

Across all cohorts, the weighted average liability scale variance in SI explained by the SI PRS was 0.49% (95% confidence interval (CI) = 0.37-0.62%) (Figure 3). Individuals in the top quintile of the SI PRS had 1.23 times higher odds (95% CI = 1.16-1.30) of SI compared with those in the middle quintile (Figure 3). The corresponding AUC was 0.54 (95% CI = 0.53-0.54). Consistent with the lower power of the discovery GWAS, the SD PRS had limited performance (Figure 3). The SB PRS explained 1.16% (95% CI = 0.90-1.44%) of the variance in SB and individuals in the top PRS quintile had 1.43 times higher odds (95% CI = 1.29-1.62) of SB than those in the middle quintile (Figure 3). The corresponding AUC was 0.57 (95% CI = 0.56-0.58). The SB PRS also explained 0.90% (95% CI = 0.69-1.11%) of the variance in SD (Figure 3). The performance of each suicidality PRS was generally higher in EUR target cohorts than in other cohorts, consistent with the more limited power of the AFR, EAS, and LA discovery GWAS to contribute to the multi-ancestry PRS (Figure 3).

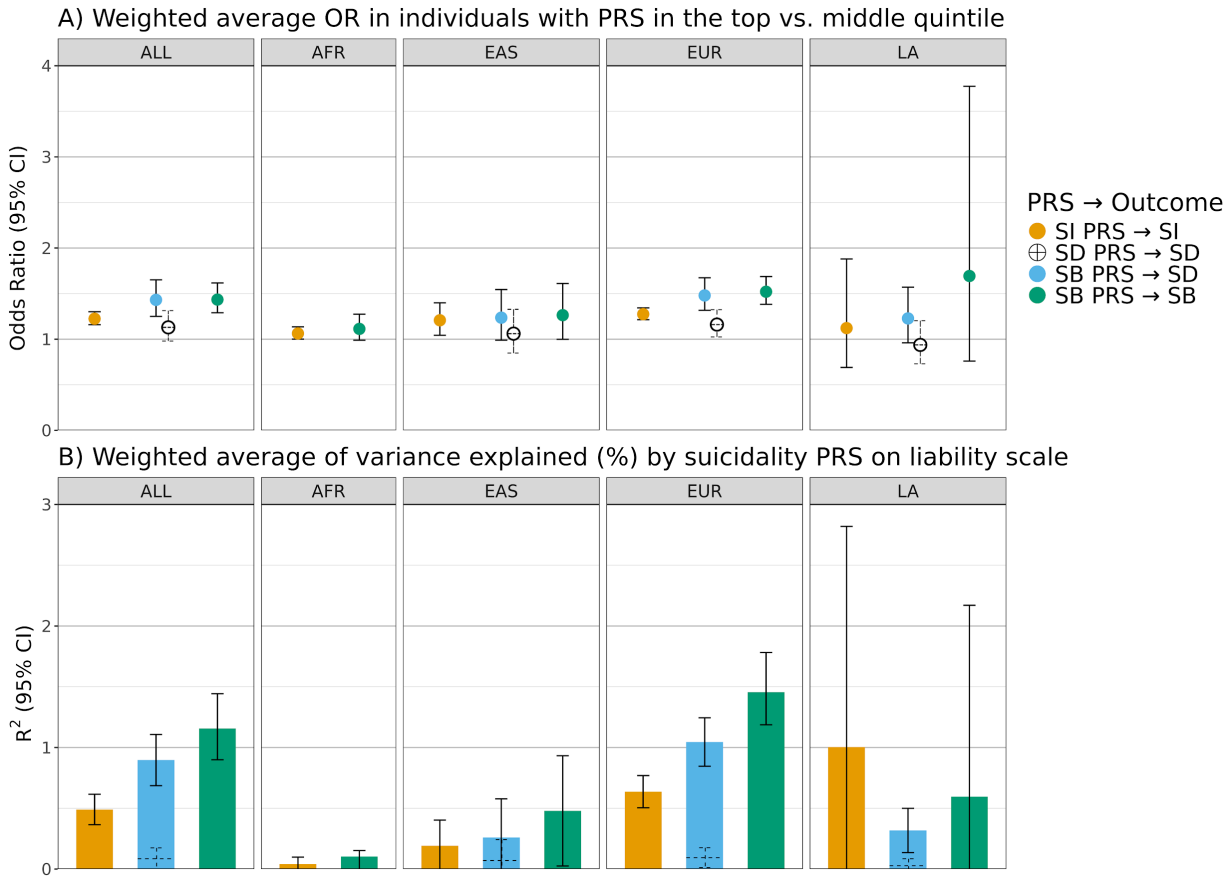

Figure 3. Polygenic risk score (PRS) prediction of suicidality phenotypes across genetic ancestries/populations. A) Effective sample size N_{eff} weighted average odds ratios (OR) with 95% confidence intervals (CIs) for the association between suicidality PRS and the corresponding outcome phenotype. ORs compare individuals in the top PRS quintile to those in the middle quintile, stratified by genetic ancestry/population group: all cohorts combined (ALL), African (AFR), East Asian (EAS), European (EUR), and Latin American (LA). **B)** N_{eff} weighted average variance explained (R^2, %) by suicidality PRS on the liability scale, with 95% CIs, across the same genetic ancestry/population groups described above. Colors indicate PRS-phenotype associations: suicidal behavior (SB) PRS predicting SB in green, SB PRS predicting suicide death (SD) in blue, and suicidal ideation (SI) PRS predicting SI in yellow. Cohort-specific results are provided in Supplementary Tables 18-21 and Supplementary Figures 17-19.

Fine-mapping and gene prioritization implicate putative causal SNPs and credible genes for SI and SB

Fine-mapping was performed on GWS loci for SI, SD (EUR) and SB using GWAS summary statistics by applying complementary methods: SuSiE[27] (statistical fine-mapping), PolyFun+SuSiE[28] (functional fine-mapping), and SuSiEx[29] (multi-ancestry fine-mapping). At a

posterior inclusion probability (PIP) threshold >0.5, we identified nine putative causal SNPs for SI, none for SD, and 18 for SB (Supplementary Table 22). These SNPs were mapped to 14 genes using Variant Effect Predictor, based on their genomic position relative to annotated transcripts and regulatory features (Figure 4).

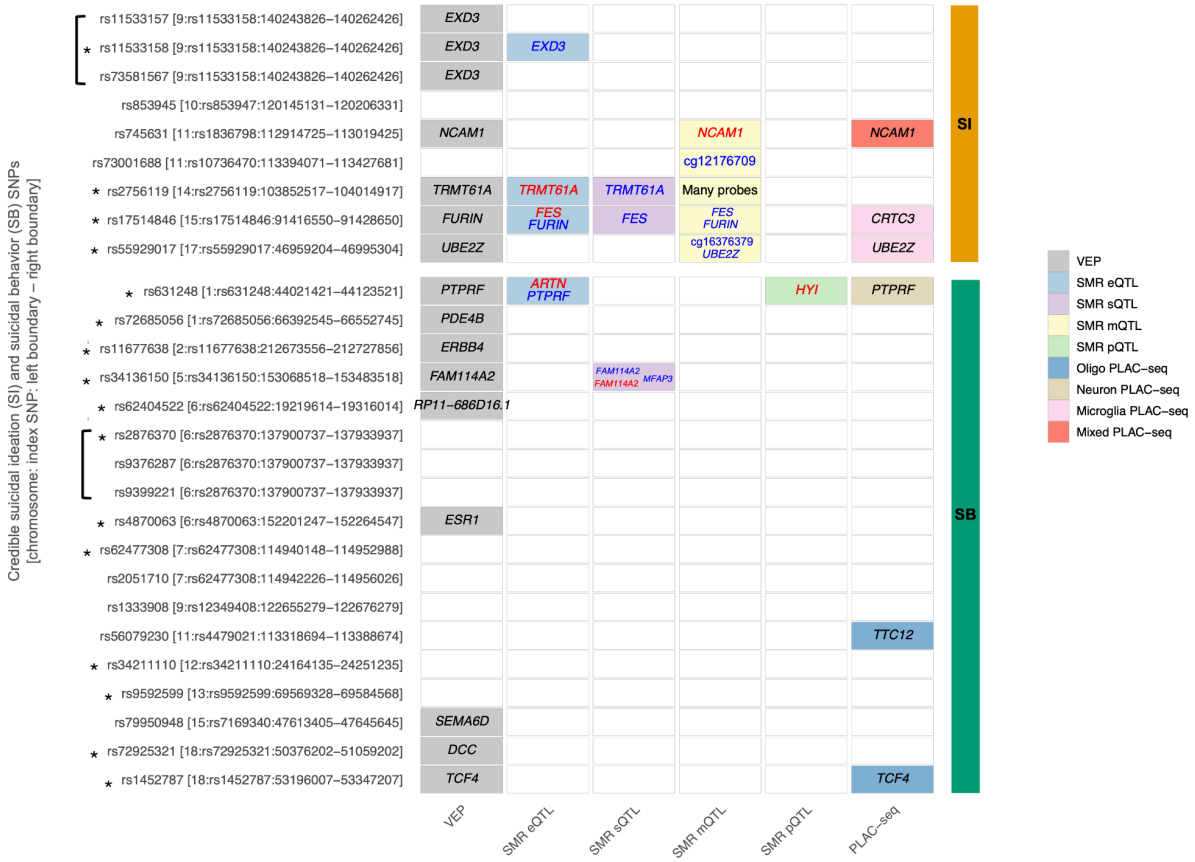

Figure 4. Gene prioritization results for suicidal ideation and suicidal behavior. The y-axis displays credible SNPs for suicidal ideation (SI) and suicidal behavior (SB), followed by the genome-wide significant locus labelled as [chromosome number:index SNP:left range: right range], with ranges denoting the fine-mapping windows. The asterisk symbol (*) indicates credible SNPs that are also the index SNP in the locus. Multiple credible SNPs from the same locus are enclosed in a square bracket. The x-axis displays the results from various gene prioritization analyses, including VEP annotations, SMR analyses of brain eQTL, sQTL, mQTL and pQTLs, and the integration of PLAC-seq data. Methylation probe IDs were mapped to genes using the EWAS atlas. For SMR analyses, gene and probe names are color coded as blue or red based on the effect direction of the risk allele on the molecular feature. The label 'Many probes' is assigned to credible SNPs for which multiple methylation probes are implicated via SMR analyses. The label 'Mixed PLAC-seq' indicates cases where evidence for a SNP came from more than one PLAC-seq cell type (Neuron and Oligo PLAC-seq).

Additionally, we applied an integrative quantitative trait locus (QTL)-based strategy, leveraging fine-mapped SNPs for SI and SB as instruments in Summary data-based Mendelian

randomization (SMR) analyses[30], including brain QTLs for expression (eQTLs)[31], splicing (sQTLs)[31], methylation (mQTLs)[32] and protein abundance (pQTLs)[33]. This approach identified 11 credible genes (Figure 4, Supplementary Table 23). We also mapped SNPs to target genes using PLAC-seq chromatin interaction data from brain cell-types[34], which additionally prioritized *TTC12* and *CRTC3* (Figure 4). The credible genes identified have roles in neurodevelopmental and signaling processes, among other functions (Supplementary Table 24). We investigated whether the protein products of credible genes are drug targets using the Drug-Gene Interaction Database v4[35] and Connectivity Map[36]. For SI, we found several drug-gene pairings, including *NCAM1*, a target of the serotonin-norepinephrine reuptake inhibitor duloxetine, *KLC1 and FES*, both targets of cancer therapies, and *FURIN*, which is targeted by the anti-inflammatory drug pirfenidone (Supplementary Table 25). For SB, we found three main drug targets: *PDE4B* (an anti-inflammatory target), *ERBB4* (a tyrosine kinase receptor often targeted by cancer drugs), and *ESR1* (an estrogen receptor target) (Supplementary Table 25).

Enrichment analyses reveal neurobiological underpinnings of suicidality

Gene set enrichment analyses were conducted using MAGMA[37] and GSA-MiXeR[38] to identify biological pathways associated with suicidality phenotypes. No gene sets reached significance for SI or SD. MAGMA analysis of SB identified six significantly enriched gene sets, primarily pointing to synaptic and neuronal structure and function (Supplementary Table 26). GSA-MiXeR results indicated 13 gene sets enriched for SB (Akaike information criterion (AIC) > 0; Supplementary Table 27), although leave-one-gene-out analyses revealed that almost all were driven by single, highly enriched genes, suggesting limited contributions from the remaining genes in the set (Supplementary Table 27).

MAGMA and GSA-MiXeR did not identify any of the same gene sets, with MAGMA-implicated gene sets being larger and broader and GSA-MiXeR-implicated gene sets being smaller and more specific, consistent with prior studies[38]. However, there was functional convergence in the gene sets implicated between the methods; for instance, the dopamine

neurotransmitter receptor activity gene set identified by GSA-MiXeR overlaps with the larger postsynapse, synaptic membrane, and postsynaptic membrane gene sets identified by MAGMA.

We also performed tissue-specific gene expression enrichment analyses using MAGMA via FUMA[39]. Among 30 general GTEx tissue types[40], significant enrichment was observed for SI and SB in brain and pituitary tissues, while SD showed enrichment only in brain tissue (all $p < 0.00167$) (Supplementary Table 28). Similar analysis of 54 specific GTEx tissue types revealed widespread enrichment across brain regions (Supplementary Table 29).

To identify cell-type specific enrichment by brain region, we performed stratified LDSC[41] to estimate h^2_{SNP} enrichment in the top decile of expression proportion (TDEP) genes across 31 brain cell type superclusters and 461 brain cell type clusters, as previously described[42,43]. SI showed significant heritability enrichment in four neuronal cell-type superclusters: amygdala excitatory, medium spiny neuron, miscellaneous, and thalamic excitatory (Figure 5B, Supplementary Table 30). No cell-type superclusters showed significant enrichment in the analysis of SD. For SB, four neuronal superclusters were enriched: amygdala excitatory, caudal ganglionic eminence (CGE) interneuron, eccentric medium spiny neuron, and hippocampal CA1-3 (Figure 5, Supplementary Table 30).

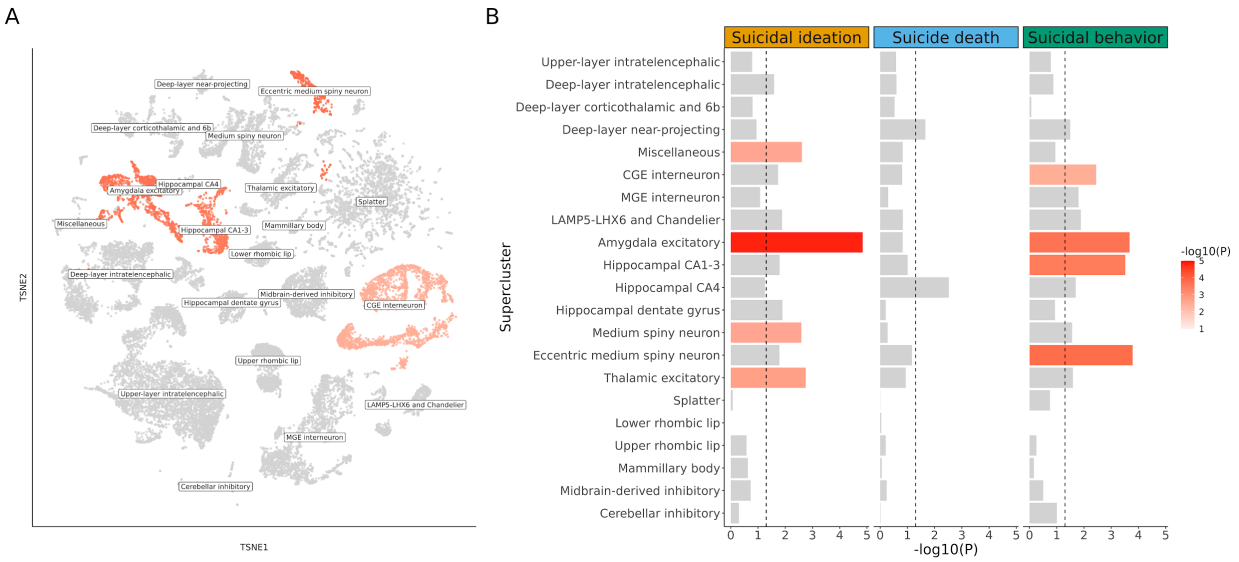

Figure 5: Supercluster-level h^2_{SNP} enrichment for suicidality phenotypes. A) t-Distributed Stochastic Neighbour Embedding (tSNE) plot for suicidal behavior (naming convention and source of the single-cell data from Siletti et al.[43]) colored by -log10(P) enrichment values. Grey indicates non-significantly enriched superclusters (false discovery rate > 0.05). **B)** Bar plot showing enrichment significance for 21 neuronal superclusters for suicidal ideation (yellow), suicide death (blue), and suicidal behavior (green). For all phenotypes, no enrichment was observed in non-neuronal cell clusters (Supplementary Table 30).

DISCUSSION

This study represents the largest and most ancestrally diverse GWAS of suicidality phenotypes to date. We identified 77 unique loci, including 59 previously unreported, representing a marked expansion of the known genetic architecture for suicidality. Consistent with prior literature[17,20], h^2_{SNP} was highest for SA (6.6%) and SB (6.7%). We also observed significant h^2_{SNP} for SI (3.0%) and SD (5.1%), reinforcing their polygenic nature. Notably, h^2_{SNP} estimates were not significant in the AFR, EAS, or LA GWAS, and low mean χ^2 estimates reflect persistent disparities in power, likely due to sample size imbalances across these groups and the limited performance of existing reference panels for admixed populations.

Our results demonstrate both shared and distinct genetic etiologies across suicidality phenotypes. Many GWS loci overlapped between the SA and SB meta-analyses (Figure 1E) and indeed they were based on mostly similar component GWAS. Nonetheless, the SB GWAS uncovered 22 additional loci not identified in SA, highlighting the potential for additional discovery through larger sample sizes[44]. The largest previous SB GWAS identified 12 GWS loci[20], of which nine were GWS, two were near GWS ($p < 1.35 \times 10^{-6}$), and one had a minimum p-value of 0.005 in the current European or multi-ancestry SB GWAS. The SI GWAS yielded fewer GWS loci, but still revealed partial overlap with loci implicated in SA and SB. In contrast, we identified only two GWS loci in the EUR SD GWAS, neither of which has been reported in previous suicidality GWAS, nor emerged in the current GWAS of SI, SA, or SB. While a previous SD GWAS[21] identified six variants across two loci on chromosomes 13 and 15, these loci were not associated with SD in the current study (p>0.01). This discrepancy may reflect greater

heterogeneity introduced by combining multiple cohorts in meta-analysis, and heterogeneity among SD cases, given recent evidence for distinct genetic profiles in those with versus without prior suicidality[45]. While the limited detection of GWS loci in the SD GWAS is likely a result of limited statistical power due to smaller case numbers, the specificity of these loci may also suggest distinct genetic risk factors for SD.

Further characterizing the genetic architecture of suicidality, we observed strong genetic correlations among SI, SD, and SB, although all were significantly below 1, supporting prior evidence that these phenotypes, while related, are genetically distinguishable and may reflect different stages or pathways in the suicidality continuum[9,46–48]. Notably, while the genetic correlations of SD with SI and SB were high, both were lower than the correlation between SI and SB, reinforcing the interpretation of genetic specificity in SD. These findings also indicate that, although SI and SB share some genetic risk factors, their biological etiology may be partially distinct, as evidenced by differences in h^2_{SNP} and GWS loci between them, along with the genetic correlation significantly < 1. For SI, the GWAS results may additionally be influenced by the complexity and heterogeneity of the phenotype, which represents an internal cognitive process that ranges in severity and duration.

Complementary fine-mapping and gene prioritization approaches identified nine and 18 putative causal SNPs and prioritized seven and 13 credible genes for SI and SB, respectively. None of the GWS loci that overlapped between SI and SB were successfully fine-mapped for both, and thus we could not determine if the two phenotypes share putative causal SNPs or credible genes. Ten prioritized genes had been previously associated with suicidality phenotypes[17,19,20,23], for example *ESR1* and *FURIN*. The ten previously unreported genes included *CRTC3*, *NCAM1*, *TRMT61A*, and *UBE2Z* for SI and *ERBB4*, *FAM114A2*, *HYI*, *MFAP3*, *PTPRF,* and *TCF4* for SB, several of which have been associated with other psychiatric and behavioral phenotypes (Supplementary Table 24). For example, *NCAM1*, *PTPRF*, and *TCF4* were identified in GWAS of impulsivity-related phenotypes[49], and *NCAM1* and *EXD3* have been implicated in externalizing behavior[50,51] and borderline personality disorder[52], respectively.

Additionally, *ESR1* and *FURIN* were credibly associated with bipolar disorder and MDD in similar gene prioritization analyses[53,54].

To further contextualize the biological underpinnings of suicidality, we conducted enrichment analyses spanning gene set, tissue and cell-type expression. Collectively, these analyses converged on neuronal pathways and brain-expressed genes as key contributors to suicidality phenotypes. Gene set enrichment analyses using MAGMA implicated broad synaptic gene sets (consistent with interactome evidence[55]) in SB, while GSA-MiXeR highlighted more targeted processes, such as dopaminergic signaling. Many of the GSA-MiXeR-results appeared to be driven by a small number of highly associated genes, and should be interpreted with caution, as the enrichment may reflect a single gene rather than involvement of the entire pathway.

SI and SB both showed significant enrichment in genes highly expressed in pituitary and brain tissues, and more granular analyses of brain cell-types revealed heritability enrichment across multiple neuronal populations. Enrichment was absent in non-neuronal superclusters, consistent with results for other psychiatric phenotypes[42]. For example, SI showed enrichment in the amygdala excitatory, medium spiny neuron, and thalamic excitatory superclusters, while SB also showed enrichment in amygdala excitatory, along with the CGE interneuron, eccentric medium spiny neuron, and hippocampal CA1-3 superclusters. Excitatory and medium spiny neurons are closely aligned with glutamatergic and dopaminergic signaling pathways, respectively, which have long been central to psychopharmacology[56,57]. The hippocampal CA1-3 enrichment may additionally point toward involvement of neuronal populations critical for memory formation. Memory processes may be especially relevant to suicide risk; epigenetic alterations in the hippocampus of suicide decedents have been linked to disrupted expression of genes involved in memory[58], which may also impair processes like problem-solving and stress regulation that protect against suicidal behavior[59].

Prior work has shown that many of the neuronal populations implicated here are concentrated in specific brain regions[42], such that cell-type enrichment may also reflect brain

region involvement. Most of the enriched cell types are primarily found in subcortical structures, such as the limbic system (i.e., hippocampus, amygdala, and hypothalamus), basal forebrain, thalamus, and midbrain[42]. Suicidality phenotypes showed a relative absence of enrichment in cell type superclusters primarily located in the cortex, in contrast to other psychiatric phenotypes[42]. This may reflect distinct neurobiological mechanisms underlying suicidality and potential dysregulation in processes more closely tied to subcortical circuitry, such as impulsivity, emotion processing, stress reactivity, and threat detection, which may play critical roles in suicidal crises[60–62].

Polygenic prediction of SI has historically been limited, with SI PRS explaining just 0.2% of phenotypic variance[19]. Here, using multi-ancestry PRS derived from our better-powered GWAS, we observe modest improvement, with SI PRS explaining ~0.5% of the variance (Figure 3). PRS generally performed better for SB, with an average R^2 of ~1.2% and a maximum of 2.9% in a single cohort (previous maximum R^2 ~1.1%)[20]. These improvements likely reflect larger discovery GWAS and the advantages of multi-ancestry PRS methods[26]. In contrast, the weaker performance of the SD PRS is likely due to smaller discovery GWAS sizes and limited ancestral diversity. Notably, these PRS alone are not adequate for individual-level clinical prediction or risk stratification. However, with future advances, they may provide clinical utility when combined with complementary measures of risk derived from electronic health records[63] and other data modalities. At present, the greatest utility of PRS lies in research to identify genetic contributions to specific facets of suicidality[45,64], or examine associations between genetics and other risk factors[65,66] or changes in brain morphology[67].

We also examined genetic correlations between suicidality phenotypes and a broad set of psychiatric and behavioral traits. Compared to the genetic correlations between SI, SD, and SB, we observed even lower correlations with non-suicidal self-injury, pointing to a genetic component potentially unique to suicidal intent. Consistent with previous evidence that internalizing disorders are key contributors to suicidality at both the phenotypic and genetic levels[8,17,68,69], MDD, anxiety, and PTSD showed some of the strongest genetic correlations with

suicidality. SB, in particular, showed stronger correlations with externalizing behavior[50,51] and associated traits, such as problematic substance use and ADHD, than SI. This is consistent with the role of behavioral disinhibition in the transition from ideation to action[70–73]. We did not exclude cases or controls based on psychiatric diagnoses to avoid bias from comorbidities. While this approach maximizes inclusivity and power, it may also inflate the observed genetic correlations with psychiatric disorders. Still, no genetic correlations with psychiatric disorders suggested total genetic sharing, reinforcing the view that genetic risk underlying suicidality cannot be fully explained by genetic liability to psychiatric disorders[17,74].

Several other limitations of this study should be acknowledged. First, we analyzed SI, SA, SD, and SB as separate phenotypes, despite their shared features and the increased power that analyzing these phenotypes together might provide. However, this was a deliberate design choice aimed at generating high-quality, specific GWAS that will provide a valuable resource to the broader psychiatric genetics community[5]. Future work from PGC SUI will involve meta-analytic and multivariate methods, like genomic structural equation modeling[75], to formally test the extent of genetic overlap and specificity both across suicidality phenotypes and between suicidality phenotypes and psychiatric disorders. Second, we were unable to perform “nested” GWAS and thus cannot fully disentangle the genetic contributions to progression between phenotypes. For example, a GWAS of SB restricted to individuals with SI, which would enable comparisons between those who experience SI without acting on it, and those who progress to action. Similarly, we also could not parse potential differences in the pathways that lead to SA versus SD. Finally, although fine-mapping and enrichment analyses identified credible genes associated with suicidality risk and promising biological pathways, these findings should be investigated in functional experiments to validate mechanisms.

In conclusion, this study represents the largest and most ancestrally diverse GWAS of suicidality phenotypes to date. Fine-mapping and enrichment analyses highlight neuron-specific expression patterns and implicate subcortical brain regions and synaptic signaling pathways in the etiology of suicidality. Differences in heritability, genetic correlation, and predictive

performance across SI, SA, SD and SB support both shared and phenotype-specific genetic influences. These findings provide a foundational resource for future functional, translational, and cross-trait studies aimed at understanding the biological factors underlying suicide risk. Ultimately, biological insights represent just one component of suicidality risk, and their integration with clinical and environmental factors will be critical for informing and developing prevention efforts.

METHODS

Cohorts and phenotype definitions

For each cohort, we provide detailed information on sample size, genetic ancestry/population, ascertainment methods, phenotyping approach, type of data received, and ethical approvals summarized in Supplementary Table 1 and described further in the Supplementary Note. To support harmonization across cohorts, each study followed a previously published phenotyping protocol by PGC SUI[5] to guide the construction of case and control definitions. We classified suicidality phenotypes into three primary categories based on standard definitions[11]: 1) SI, wherein an individual contemplates ending their own life with or without a specific plan; 2) SA, wherein an individual takes action to cause harm to themselves with at least some intent to die; 3) SD, or death caused by intentional action to take one's own life.

The SI GWAS included SI cases (some of whom may have also reported SA) and controls without evidence of SI, SA, or SD. The SA GWAS included only non-fatal SA cases and controls who did not have evidence of SA or SD, but who may have reported SI. Our definition of SA required at least some intent to die, such that non-suicidal self-injury was not classified as SA. The SD GWAS included SD cases and controls who did not die by suicide or have evidence of SA. The SD GWAS also excluded deaths of undetermined intent, as it is unclear what

proportion of these deaths may actually be SDs[76]. In addition to these individual phenotypes, we also analyzed an aggregated phenotype, SB, defined here to include SA or SD. Thus, the SB GWAS combined SA and SD cases and used controls without evidence of either. Psychiatric diagnoses were not used as exclusion criteria for cases or controls in any GWAS to avoid bias related to comorbid psychiatric conditions. SI and SA data were obtained through a variety of sources, including structured psychiatric interviews, self-report questionnaires, and International Classification of Diseases (ICD) codes, or combinations of these sources (Supplementary Table 1). SDs were identified using cause of death codes, Coroners' reports, or Medical Examiner's reports (Supplementary Table 1). Further details on specific case/control criteria and the rationale behind these decisions can be found in the PGC SUI phenotyping protocol[5].

If a study included individuals from multiple genetic ancestries or populations, individuals were grouped according to the ancestry or population they were most genetically similar to, and here are referred to as distinct analytic cohorts. Groups were determined in each cohort using principal component analysis and reflects which population (from the reference panel used by each cohort) an individual was most genetically similar to. In total, the analyses included 37 SI cohorts (N_{Cases} = 259,747 and $N_{Controls}$ = 1,309,943), 46 SA cohorts (N_{Cases} = 64,993 and $N_{Controls}$ = 1,269,037), 49 SB cohorts (N_{Cases} = 75,300 and $N_{Controls}$ = 1,311,895), and 7 SD cohorts (N_{Cases} = 9,197 and $N_{Controls}$ = 668,162) (Table 2). The current SB meta-analysis represents an increase of 31,492 cases not included in the previous GWAS meta-analysis of SB[20].

Genotyping, quality control, imputation, and analysis

Studies either securely transferred individual-level genotype and phenotype data to the PGC server for central analysis or shared GWAS summary statistic data (Supplementary Table 1). All samples underwent standard genotyping performed by the collaborating research teams, the details of which are provided in the Supplementary Note. Where individual-level genotype data were provided, quality control (QC) and imputation were performed using the Rapid Imputation for COnsortias PIpeLIne (RICOPILI) developed by the PGC and described in detail

previously[77]. In summary, QC parameters for retaining SNPs and subjects were: SNP missingness < 0.05 (before sample removal), subject missingness < 0.02, autosomal heterozygosity deviation (F_{het}) < 0.2, SNP missingness < 0.02 (after sample removal), difference in SNP missingness between cases and controls < 0.02 and SNP Hardy-Weinberg equilibrium ($P > 10^{-10}$ in cases, $P > 10^{-6}$ in controls). When sample overlap or relatedness was suspected, relatedness was calculated within or across cohorts using identity by descent, and one of each pair of related individuals (pi_hat > 0.2) was excluded, with a preference for retaining cases. Genotype imputation was performed using the prephasing/imputation stepwise approach implemented in Eagle (v2.3.5; https://alkesgroup.broadinstitute.org/Eagle/)[78] and Minimac3 (v2.0.1; https://genome.sph.umich.edu/wiki/Minimac3)[79] to the Haplotype Reference Consortium (HRC) reference panel (v1.1)[80]. GWAS were then performed separately in each cohort and genetic ancestry or population group using PLINK 1.9 (v1.90b6.4)[81] by comparing imputed marker dosages under an additive logistic regression model between cases and controls, covarying for the first five ancestry principal components (PCs) and only additional PCs if they were associated with case-control status. Studies contributing GWAS summary statistics performed QC, imputation, and GWAS locally using comparable procedures (see the Supplementary Note for details). Studies submitting GWAS summary statistics sometimes strayed from the standard analytic procedure to account for unique components of their study design. For example, some family-based cohorts used GWAS methods that account for relatedness within the model, such that related individuals were retained and analyzed in the sample (see the Supplementary Note for details).

GWAS quality control and meta-analysis

GWAS summary statistics from all cohorts underwent standard QC procedures and data cleaning after being transferred to the PGC SUI analytic team. First, SNPs were filtered from the GWAS summary statistics from each cohort using sample minor allele frequency (MAF) < 1% and sample MAF corresponding to a minor allele count (MAC) < 10 in cases or controls

(whichever had smaller N), in order to control test statistic inflation at low MAFs from small cohorts. We then checked summary statistics files for missingness, duplicate SNPs, correct build (hg19), and test statistic values being reported on the appropriate scale. Any errors or inconsistencies found were remedied before moving on to allele frequency checks. SNP allele frequencies from the GWAS summary statistics were compared against allele frequencies from the (appropriate ancestry/population-specific) 1000 Genomes Project (1KG)[82] and HRC reference panels[80]. If a SNP had an allele frequency difference > 0.20 between the summary statistics and both of the reference panels, it was removed. Finally, we used DENTIST[83] to identify and remove SNPs with errors in the test statistics or heterogeneity between the GWAS and reference.

After each individual GWAS passed cleaning and QC, we conducted multi-ancestry meta-analyses of each phenotype using an inverse variance-weighted fixed-effects model in METAL[84], implemented using RICOPILI[77]. We also conducted ancestry/population-specific meta-analyses when GWAS summary statistics for a specific ancestry/population and phenotype were available from more than one cohort. We defined GWS loci as regions surrounding SNPs with p-values $< 5 \times 10^{-8}$ with LD $r^2 > 0.1$, within a 3,000 kb window based on the LD structure from the HRC reference panel, using the EUR panel for multi-ancestry meta-analyses and the ancestry/population-matched reference panel for ancestry/population-specific meta-analyses.

To determine if a GWS locus identified in the current analysis was previously associated with suicidality, or was overlapping amongst suicidality phenotypes, we used the intersect function from bedtools (v2.31.1)[85] to check for overlapping regions. To determine which GWS loci in the current analysis were previously unreported, we checked for overlap with GWS loci identified in seven different previous GWAS of various suicidality phenotypes[17–21,23,24].

For all post-GWAS analyses, unless otherwise specified, we used the multi-ancestry GWAS of SB, SI, and SD with a filter applied to remove SNPs with a per-SNP effective sample size (N_{eff}) more than 3 standard deviations below the mean per-SNP N_{eff}. For follow-up analyses

requiring an LD reference panel, we used a EUR LD reference panel with multi-ancestry GWAS. To assess the appropriateness of this reference panel, we followed the approach outlined by Yengo et al. (2022)[86], which involves evaluating various population stratification metrics. Specifically, we examined heterogeneity (Cochran's Q-statistic), the fixation index of genetic differentiation (F_{ST}) across SNPs[87,88], and the LDSC attenuation ratio[22]. For all phenotypes, the mean Q-statistic across SNPs in the multi-ancestry GWAS indicated only moderate heterogeneity across ancestry and population groups (5-38%) (Supplementary Table 31). The mean F_{ST} between the EUR and multi-ancestry GWAS SNPs was low for all phenotypes (0.0003-0.003), consistent with values observed between populations of the same genetic ancestry[86]. Finally, all LDSC attenuation ratios calculated using the 1KG EUR LD scores were below 20%, indicating a good match between the EUR LD reference and the multi-ancestry GWAS[22] (Supplementary Table 31). Taken together, these results supported the use of the EUR reference as an approximation for analyses involving the multi-ancestry GWAS.

Linkage Disequilibrium Score Regression

We used LDSC[22] to estimate the h^2_{SNP} of SI, SA, SD, and SB from ancestry/population-specific and multi-ancestry GWAS using ancestry matched LD scores (Supplementary Table 15). We converted h^2_{SNP} to the liability scale assuming lifetime prevalence of 0.2% for SD[21], 2% for SA/SB (median of the range reported worldwide)[89], and 9% for SI[89]. We acknowledge that prevalence may vary across populations, and as such h^2_{SNP} estimates on the liability scale could be over- or underestimated, although the significance of the estimate is unaffected. We also used LDSC to estimate genome-wide genetic correlations[25] between the suicidality phenotypes and other phenotypes previously shown to have genetic relationships with suicidality (described in detail in Supplementary Table 15). For the non-suicidality phenotypes, we used multi-ancestry GWAS when available and when the LDSC statistics indicated a good match to the EUR LD reference panel (as described above). For all other non-suicidality phenotypes, we used European ancestry GWAS. The genetic ancestry of the

non-suicidality phenotypes used are provided in Supplementary Table 15. To account for multiple testing, we applied the Benjamini-Hochberg procedure, considering correlations with FDR $p < 0.05$ as statistically significant; this threshold was applied as the maximum acceptable FDR across all genetic correlation analyses. To test for a significant difference between two genetic correlations, we followed the method from Coleman et al.[90] which applies a block jackknife.

Polygenic Risk Scoring

We used PRS-CSx (version from Nov 21, 2024)[26] to compute multi-ancestry meta-analyzed posterior effect size estimates and generate multi-ancestry PRS for suicidality phenotypes in 25 target cohorts. For each target cohort, we used a leave-one-out approach, excluding that cohort from the discovery GWAS used to calculate the PRS. PRS-CSx jointly models ancestry/population-specific GWAS summary statistics using a shared continuous shrinkage prior, enabling correlated yet varying effect sizes across ancestries/populations. Thus, for discovery GWAS we used ancestry/population-specific GWAS for as many genetic ancestry/population groups as available for each phenotype. PRS-CSx improves effect size estimation through sharing information between GWAS, leveraging LD diversity, and accounting for population-specific allele frequencies. We used the "--auto" option to automatically infer model parameters, and then combined population-specific posterior effect size estimates using the"--meta" option[26]. Associations between PRS and suicidality phenotypes were tested separately within each cohort using logistic regression (via the glm() function in R[91]), with the PRS standardized (mean = 0, standard deviation = 1) and the same covariates used in that cohort's GWAS. We converted the variance explained by the PRS (R^2) to the liability scale using the method from Lee et al. (2012)[92] to account for the proportion of cases in the target cohort and the population prevalences from above. We calculated 95% CIs for the liability scale R^2 using non-parametric bootstrapping (10,000 resamples) via the boot package in R[93,94]. We also calculated ORs comparing individuals in the top quintile/decile of the PRS distribution to those in

the middle quintile/decile, using standardized residuals of the PRS (adjusted for the same covariates used in the logistic regression). We calculated AUC statistics using the pROC package in R[95] with models including only the scaled PRS as a predictor following standard PGC procedures[77]. To ensure consistency across cohorts, PRS weights were centrally generated on the PGC server and subsequently distributed to study analysts along with standardized analysis scripts (https://github.com/sarahcolbert/PGC_SUI_PRS). Finally, the R^2s, ORs, AUCs, and their CIs were meta-analyzed using effective sample size weighting to calculate N_{eff}-weighted averages within and across genetic ancestries/populations.

Fine-mapping

We conducted statistical and functional fine-mapping of GWS loci for SI, SB, and SD (EUR). The fine-mapping windows were defined as regions surrounding SNPs with p-values < 5×10^{-8} with LD $r^2 > 0.6$, within a 3,000 kb window based on the LD structure from the HRC EUR reference panel (v1.1)[80]; and loci in the MHC (chromosome 6: 28,000,000–34,000,000, GRCh37) were excluded due to the complexity of the LD structure in this region. After these filters, we fine-mapped 12 GWS loci from SI multi-ancestry, 51 loci from SB multi-ancestry, and two loci from SD EUR. We employed statistical fine-mapping using SuSiE[27] and functional fine-mapping with Polyfun+SuSiE[28], both implemented via the SAFFARI pipeline[96], to the multi-ancestry SI and SB results and to the EUR ancestry SD results. LD estimates were obtained from the HRC EUR reference panel, and the functional annotations for Polyfun+SuSiE were from the baseline-LF2.2 UK Biobank model[97]. We restricted the maximum number of causal variants per fine-mapped region to the default of five (K=5)[27], and performed single variant fine-mapping (no LD), assuming one causal variant in each locus (K=1), as this approach does not require population-accurate estimation of LD. Additionally, we applied SuSiEx[29], a multi-ancestry fine-mapping method, leveraging ancestry/population-specific meta-analysis results and the corresponding ancestry/population LD reference panel from the HRC. For both SI and SB, the SuSiEx analyses included summary statistics from EUR, LA,

AFR, EAS, and SAS GWAS. Putative causal SNPs were defined as those with PIP > 0.5 and included in the 95% credible set for the multivariant analyses (K=5) and as those with PIP > 0.5 for the single variant analysis.

Gene prioritization analyses

Fine-mapped SNPs were mapped to genes through variant annotation using the Ensemble Variant Effect Predictor (VEP; GRCh37, Ensembl release 114, https://www.ensembl.org/info/docs/tools/vep/index.html). Furthermore, we integrated brain quantitative trait loci (QTL) data to link fine-mapped variants to molecular features and elucidate potential molecular mechanisms by which they increase risk of SB or SI. Data on eQTLs and sQTLs were obtained from the BrainMeta study (version 2), which comprised RNA-sequencing data of 2,865 brain cortex samples from 2,443 unrelated individuals of EUR ancestry with genome-wide SNP data[31]. Data on mQTLs were obtained from the BrainMeta study[32], a meta-analysis of adult cortex or fetal brain samples, comprising 1,160 individuals with methylation levels measured using the Illumina HumanMethylation450K array. We also used pQTLs from a study of brain proteomes of 1,277 individuals of EUR ancestry primarily from the dorsolateral prefrontal cortex[33]. Summary data-based Mendelian randomization (SMR; v1.3; https://yanglab.westlake.edu.cn/software/smr/)[30] with subsequent heterogeneity in dependent instruments (HEIDI) tests was performed for each QTL dataset using the fine-mapped SNPs as the QTL instruments. Using the multi-ancestry SB and SI GWAS and QTL summary statistics, each putative causal SNP was analysed as the target SNP for probes within a 2-Mb window on either side using the --extract-target-snp-probe option in SMR. The EUR HRC LD reference panel was used for the analyses of the multi-ancestry meta-analysis. SNP-QTL probe combinations were filtered according to the following criteria: $P_{QTL} < 5x10^{-8}$ and the significance threshold for the HEIDI test was $P_{HEIDI} \geq 0.01$. To improve the robustness of our SMR findings, we repeated our SMR analysis using the European GWAS meta-analyses for SI and SB, while

keeping the remaining inputs and filtering criteria unaltered. This analysis further solidified our findings while showing cross-ancestry generalizability (Supplementary Table 23).

Furthermore, we annotated SNPs to genes using PLAC-seq data from brain cell-types[34], providing an additional layer of chromatin interaction-based evidence. Specifically, genes were annotated when either end of a chromatin interaction loop was linked to them, using previously published epigenomic data from purified bulk H3K27ac and H3K4me3 ChIP–seq of brain cell-types[34].

Enrichment analyses

Gene, gene set, and tissue enrichment analyses

We conducted gene, gene set and tissue-level enrichment analyses to explore the biological underpinnings of suicidality phenotypes. Gene and gene set analyses were conducted following previously established methods[37,38]. Briefly, associations between each suicidality phenotype and genes or gene sets were assessed using a SNP-wise mean model with a ± 10-kb window implemented in MAGMA[37]. Gene sets were obtained from Molecular Signature Database (MsigDB) v7.0[98] and 7,167 gene sets were tested after removing those with < 10 genes. Furthermore, we also used GSA-MiXeR[38] to assess gene set enrichment for 10,475 gene sets from the Gene Ontology[99] and SynGO[100] databases. A positive AIC value indicates that the full MiXeR model (with gene-specific contributions to the heritability) has a better fit than the baseline model (in which individual genes do not contribute) for that gene set. Tissue enrichment analyses were also conducted using MAGMA via FUMA[39], testing for enrichment of association signals in genes expressed across 30 broad and 54 specific tissue types from GTEx v8[40]. Bonferroni correction was applied to account for multiple testing in all enrichment analyses and thresholds for specific analyses are reported in the Supplementary Tables.

Brain cell-type enrichment analysis

Following a previously published pipeline[42], we applied stratified LDSC (s-LDSC)[41] to estimate partitioned h^2_{SNP} for the TDEP genes across brain cell types. In summary, cell types were defined using the Human Brain Atlas single-nucleus RNA sequencing dataset[43], which includes over 3 million nuclei collected from post-mortem adult human brains, samples across 106 anatomical regions within 10 major brain areas. Cell types were organized into 461 clusters and 31 superclusters, for each of which we identified the top 10% of genes ranked by expression proportion. These gene sets were then used as the functional annotations in s-LDSC. Multiple testing was addressed using a FDR threshold of 5%.

Drug target analysis

After identifying plausibly associated genes (see 'Gene prioritization analyses'), we linked them to known drug targets following a previously published pipeline that demonstrated enrichment across psychiatric disorders[101]. Briefly, drug targets were obtained from gene-drug interaction database v4[35] and connectivity map[36]. We then constructed a Sankey network diagram based on genes that linked to targets of specific pharmaceuticals. The Sankey network is a two dimensional graph network where drugs that relate to more discovered genes or vice versa are shown to have larger edges in the overall model. Those edges are then plotted in two dimensions with the connections between gene and drug visualized as a graph. It should be noted that this approach does not evaluate whether the actions of these agents align with or counteract the effects associated with genetic liability.

DATA AVAILABILITY

Summary statistics from these GWAS are available by application through the Psychiatric Genomics Consortium data access portal: (https://pgc.unc.edu/for-researchers/data-access-committee/data-access-portal/). Full summary statistics corresponding to the GWAS reported in this study are available for the SI and SA GWAS. For the SB and SD GWAS, the summary statistics available through the PGC portal exclude data from the Utah cohorts. Data from the Utah site will be available upon request from the administrative manager of the Utah Suicide Mortality Research Study

(https://rge.utah.edu/usmrs.php) and will be contingent upon the permissions required by institutional regulations and state statutes.

CODE AVAILABILITY

All software used is publicly available at the URLs or references cited. Code used to test PRS associations is available on github (https://github.com/sarahcolbert/PGC_SUI_PRS).

ACKNOWLEDGEMENTS

We thank the participants who donated their time, life experiences, and DNA to this research, and the clinical and scientific teams that worked with them.

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. 1842169 (PI Colbert) and the National Institute of Mental Health R01 MH132733 (PI Mullins). The PGC is supported by grants to UCSD (R01MH124847), UNC (R01MH124871), MGH (R01MH124851), Mount Sinai School of Medicine (R01MH124839), Cardiff University (R01MH124873), Trinity College Dublin (R01MH124875) and Washington University St. Louis (R01DA054869). The content is solely the responsibility of the authors and does not necessarily represent the official views of the US National Institutes of Health.

We thank SURF (www.surf.nl) for the support in using the National Supercomputer Snellius. Statistical analyses were carried out on the NL Genetic Cluster Computer (http://www.geneticcluster.org) hosted by SURFsara and the Mount Sinai high performance computing cluster (http://hpc.mssm.edu), which is supported by the Office of Research Infrastructure of the National Institutes of Health (Grant Nos. S10OD018522 and S10OD026880).

This research has been conducted using the UK Biobank Resource under Application Number 82087 (PI: J Coleman). This work uses data provided by patients and collected by the NHS as part of their care and support. This research was funded by the UK National Institute for Health and Care Research (NIHR) Maudsley Biomedical Research Centre (BRC). The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

Individual and cohort-specific funding acknowledgements are detailed in the Supplementary Note.

REFERENCES

1. World Health Organization. Suicide worldwide in 2021: global health estimates. https://www.who.int/publications/i/item/9789240110069 (2025).
2. World Health Organization. Preventing suicide: A global imperative. https://www.who.int/publications/i/item/9789241564779 (2014).
3. Substance Abuse and Mental Health Services Administration (SAMHSA), Center for Behavioral Health Statistics and Quality. 2023 National Survey on Drug Use and Health (NSDUH) Releases. https://www.samhsa.gov/data/data-we-collect/nsduh-national-survey-drug-use-and-health/national-releases/2023.
4. Centers for Disease Control and Prevention. Suicide Data and Statistics. *Suicide Prevention* https://www.cdc.gov/suicide/facts/data.html (2025).
5. Colbert, S. *et al.* Defining suicidality phenotypes for genetic studies: perspectives of the Psychiatric Genomics Consortium Suicide Working Group. (2025).
6. Klonsky, E. D. & May, A. M. The three-step theory (3ST): A new theory of suicide rooted in the 'ideation-to-action' framework. *Int. J. Cogn. Ther.* **8**, 114–129 (2015).
7. Beautrais, A. L. Suicides and serious suicide attempts: two populations or one? *Psychol. Med.* **31**, 837–845 (2001).
8. Franklin, J. C. *et al.* Risk factors for suicidal thoughts and behaviors: A meta-analysis of 50 years of research. *Psychol. Bull.* **143**, 187–232 (2017).
9. Colbert, S. M. C. *et al.* Distinguishing clinical and genetic risk factors for suicidal ideation and behavior in a diverse hospital population. *Transl. Psychiatry* **15**, 63 (2025).
10. Coon, H. *et al.* Absence of nonfatal suicidal behavior preceding suicide death reveals differences in clinical risks. *Psychiatry Res.* **347**, 116391 (2025).
11. Crosby, A., Ortega, L. V. & Melanson, C. *Self-Directed Violence Surveillance; Uniform Definitions and Recommended Data Elements*. https://stacks.cdc.gov/view/cdc/11997

(2011).

12. Voracek, M. & Loibl, L. M. Genetics of suicide: a systematic review of twin studies. *Wien. Klin. Wochenschr.* **119**, 463–475 (2007).
13. Nock, M. K. *et al.* Cross-national analysis of the associations among mental disorders and suicidal behavior: findings from the WHO World Mental Health Surveys. *PLoS Med.* **6**, e1000123 (2009).
14. Bachmann, S. Epidemiology of suicide and the psychiatric perspective. *Int. J. Environ. Res. Public Health* **15**, (2018).
15. Cavanagh, J. T. O., Carson, A. J., Sharpe, M. & Lawrie, S. M. Psychological autopsy studies of suicide: a systematic review. *Psychol. Med.* **33**, 395–405 (2003).
16. Fu, Q. *et al.* A twin study of genetic and environmental influences on suicidality in men. *Psychol. Med.* **32**, 11–24 (2002).
17. Mullins, N. *et al.* Dissecting the shared genetic architecture of suicide attempt, psychiatric disorders, and known risk factors. *Biol. Psychiatry* **91**, 313–327 (2022).
18. Erlangsen, A. *et al.* Genetics of suicide attempts in individuals with and without mental disorders: a population-based genome-wide association study. *Mol. Psychiatry* **25**, 2410–2421 (2020).
19. Ashley-Koch, A. E. *et al.* Genome-wide association study identifies four pan-ancestry loci for suicidal ideation in the Million Veteran Program. *PLoS Genet.* **19**, e1010623 (2023).
20. Docherty, A. R. *et al.* GWAS meta-analysis of suicide attempt: Identification of 12 genome-wide significant loci and implication of genetic risks for specific health factors. *Am. J. Psychiatry* **180**, 723–738 (2023).
21. Docherty, A. R. *et al.* Genome-wide association study of suicide death and polygenic prediction of clinical antecedents. *Am. J. Psychiatry* **177**, 917–927 (2020).
22. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
23. Kimbrel, N. A. *et al.* Identification of novel, replicable genetic risk loci for suicidal thoughts

and behaviors among US military veterans. *JAMA Psychiatry* **80**, 135–145 (2023).

24. Kimbrel, N. A. *et al.* A genome-wide association study of suicide attempts in the million veterans program identifies evidence of pan-ancestry and ancestry-specific risk loci. *Mol. Psychiatry* **27**, 2264–2272 (2022).
25. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015).
26. Ruan, Y. *et al.* Improving polygenic prediction in ancestrally diverse populations. *Nat. Genet.* **54**, 573–580 (2022).
27. Wang, G., Sarkar, A., Carbonetto, P. & Stephens, M. A simple new approach to variable selection in regression, with application to genetic fine mapping. *J. R. Stat. Soc. Series B Stat. Methodol.* **82**, 1273–1300 (2020).
28. Weissbrod, O. *et al.* Functionally informed fine-mapping and polygenic localization of complex trait heritability. *Nat. Genet.* **52**, 1355–1363 (2020).
29. Yuan, K. *et al.* Fine-mapping across diverse ancestries drives the discovery of putative causal variants underlying human complex traits and diseases. *Nat. Genet.* **56**, 1841–1850 (2024).
30. Zhu, Z. *et al.* Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.* **48**, 481–487 (2016).
31. Qi, T. *et al.* Genetic control of RNA splicing and its distinct role in complex trait variation. *Nat. Genet.* **54**, 1355–1363 (2022).
32. Qi, T. *et al.* Identifying gene targets for brain-related traits using transcriptomic and methylomic data from blood. *Nat. Commun.* **9**, 2282 (2018).
33. Wingo, A. P. *et al.* Sex differences in brain protein expression and disease. *Nat. Med.* **29**, 2224–2232 (2023).
34. Nott, A. *et al.* Brain cell type-specific enhancer-promoter interactome maps and disease-risk association. *Science* **366**, 1134–1139 (2019).
35. Freshour, S. L. *et al.* Integration of the Drug-Gene Interaction Database (DGIdb 4.0) with

open crowdsource efforts. *Nucleic Acids Res.* **49**, D1144–D1151 (2021).

36. Qu, X. A. & Rajpal, D. K. Applications of Connectivity Map in drug discovery and development. *Drug Discov. Today* **17**, 1289–1298 (2012).
37. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput. Biol.* **11**, e1004219 (2015).
38. Frei, O. *et al.* Improved functional mapping of complex trait heritability with GSA-MiXeR implicates biologically specific gene sets. *Nat. Genet.* **56**, 1310–1318 (2024).
39. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat. Commun.* **8**, 1826 (2017).
40. GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318–1330 (2020).
41. Finucane, H. K. *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* **47**, 1228–1235 (2015).
42. Yao, S. *et al.* Connecting genomic results for psychiatric disorders to human brain cell types and regions reveals convergence with functional connectivity. *Nat. Commun.* **16**, 395 (2025).
43. Siletti, K. *et al.* Transcriptomic diversity of cell types across the adult human brain. *Science* **382**, eadd7046 (2023).
44. Abdellaoui, A., Yengo, L., Verweij, K. J. H. & Visscher, P. M. 15 years of GWAS discovery: Realizing the promise. *Am. J. Hum. Genet.* **110**, 179–194 (2023).
45. Coon, H. *et al.* Different genetic liabilities to neuropsychiatric conditions in suicides with no prior suicidality. *medRxiv* 2025.05.02.25326877 (2025) doi:10.1101/2025.05.02.25326877.
46. Edwards, A. C. *et al.* On the genetic and environmental relationship between suicide attempt and death by suicide. *Am. J. Psychiatry* **178**, 1060–1069 (2021).
47. Mullins, N. *et al.* Genetic relationships between suicide attempts, suicidal ideation and major psychiatric disorders: a genome-wide association and polygenic scoring study. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **165B**, 428–437 (2014).

48. Colbert, S. M. C. *et al.* Exploring the genetic overlap of suicide-related behaviors and substance use disorders. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **186**, 445–455 (2021).

49. Sanchez-Roige, S. *et al.* CADM2 is implicated in impulsive personality and numerous other traits by genome- and phenome-wide association studies in humans and mice. *Transl. Psychiatry* **13**, 167 (2023).

50. Karlsson Linnér, R. *et al.* Multivariate analysis of 1.5 million people identifies genetic associations with traits related to self-regulation and addiction. *Nat. Neurosci.* **24**, 1367–1376 (2021).

51. Williams, C. M. *et al.* Guidelines for evaluating the comparability of down-sampled GWAS summary statistics. *Behav. Genet.* **53**, 404–415 (2023).

52. Streit, F. *et al.* Genome-wide association study of borderline personality disorder identifies 11 loci and highlights shared risk with mental and somatic disorders. *medRxiv* 2024.11.12.24316957 (2025) doi:10.1101/2024.11.12.24316957.

53. O'Connell, K. S. *et al.* Genomics yields biological and phenotypic insights into bipolar disorder. *Nature* **639**, 968–975 (2025).

54. Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium. Electronic address: andrew.mcintosh@ed.ac.uk & Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium. Trans-ancestry genome-wide study of depression identifies 697 associations implicating cell types and pharmacotherapies. *Cell* **188**, 640–652.e9 (2025).

55. Sokolowski, M. & Wasserman, D. A candidate biological network formed by genes from genomic and hypothesis-free scans of suicide. *Prev. Med.* **152**, 106604 (2021).

56. Niciu, M. J., Kelmendi, B. & Sanacora, G. Overview of glutamatergic neurotransmission in the nervous system. *Pharmacol. Biochem. Behav.* **100**, 656–664 (2012).

57. Surmeier, D. J., Ding, J., Day, M., Wang, Z. & Shen, W. D1 and D2 dopamine-receptor modulation of striatal glutamatergic signaling in striatal medium spiny neurons. *Trends*

Neurosci. **30**, 228–235 (2007).

58. Labonté, B. *et al.* Genome-wide methylation changes in the brains of suicide completers. *Am. J. Psychiatry* **170**, 511–520 (2013).
59. Richard-Devantoy, S., Berlim, M. T. & Jollant, F. Suicidal behaviour and memory: A systematic review and meta-analysis. *World J. Biol. Psychiatry* **16**, 544–566 (2015).
60. van Heeringen, K. & Mann, J. J. The neurobiology of suicide. *Lancet Psychiatry* **1**, 63–72 (2014).
61. Miller, A. B. & Eisenlohr-Moul, T. A. Biological responses to acute stress and suicide: A review and opportunities for methodological innovation. *Curr. Behav. Neurosci. Rep.* **6**, 141–150 (2019).
62. Rudd, M. D. The suicidal mode: a cognitive-behavioral model of suicidality. *Suicide Life Threat. Behav.* **30**, 18–33 (2000).
63. Detrois, K. E. *et al.* Cross-biobank generalizability and accuracy of electronic health record-based predictors compared to polygenic scores. *Nat. Genet.* 1–10 (2025).
64. Colbert, S. M. C. *et al.* Polygenic contributions to suicidal thoughts and behaviors in a sample ascertained for alcohol use disorders. *Complex Psychiatry* **9**, 11–23 (2023).
65. Johnson, E. C. *et al.* Associations between suicidal thoughts and behaviors and genetic liability for cognitive performance, depression, and risk-taking in a high-risk sample. *Complex Psychiatry* **7**, 34–44 (2021).
66. Edwards, A. C., Singh, M., Peterson, R. E., Webb, B. T. & Gentry, A. E. Associations between polygenic liability to psychopathology and non-suicidal versus suicidal self-injury. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **195**, e32982 (2024).
67. Overs, B. J. *et al.* Effects of polygenic risk for suicide attempt and risky behavior on brain structure in young people with familial risk of bipolar disorder. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **186**, 485–507 (2021).
68. Fazel, S. & Runeson, B. Suicide. *N. Engl. J. Med.* **382**, 266–274 (2020).
69. Nepon, J., Belik, S.-L., Bolton, J. & Sareen, J. The relationship between anxiety disorders

and suicide attempts: findings from the National Epidemiologic Survey on Alcohol and Related Conditions. *Depress. Anxiety* **27**, 791–798 (2010).

70. Hadzic, A. *et al.* The association of trait impulsivity and suicidal ideation and its fluctuation in the context of the Interpersonal Theory of Suicide. *Compr. Psychiatry* **98**, 152158 (2019).
71. Van Orden, K. A. *et al.* The interpersonal theory of suicide. *Psychol. Rev.* **117**, 575–600 (2010).
72. Oquendo, M. A. & Mann, J. J. The biology of impulsivity and suicidality. *Psychiatr. Clin. North Am.* **23**, 11–25 (2000).
73. Mann, J. J., Waternaux, C., Haas, G. L. & Malone, K. M. Toward a clinical model of suicidal behavior in psychiatric patients. *Am. J. Psychiatry* **156**, 181–189 (1999).
74. Mullins, N. *et al.* GWAS of suicide attempt in psychiatric disorders and association with major depression polygenic risk scores. *Am. J. Psychiatry* **176**, 651–660 (2019).
75. Grotzinger, A. D. *et al.* Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nat. Hum. Behav.* **3**, 513–525 (2019).
76. Snowdon, J. & Choi, N. G. Undercounting of suicides: Where suicide data lie hidden. *Glob. Public Health* **15**, 1894–1901 (2020).
77. Lam, M. *et al.* RICOPILI: Rapid imputation for COnsortias PIpeLIne. *Bioinformatics* **36**, 930–933 (2020).
78. Loh, P.-R. *et al.* Reference-based phasing using the Haplotype Reference Consortium panel. *Nat. Genet.* **48**, 1443–1448 (2016).
79. Das, S. *et al.* Next-generation genotype imputation service and methods. *Nat. Genet.* **48**, 1284–1287 (2016).
80. McCarthy, S. *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat. Genet.* **48**, 1279–1283 (2016).
81. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
82. 1000 Genomes Project Consortium *et al.* A map of human genome variation from

population-scale sequencing. *Nature* **467**, 1061–1073 (2010).

83. Chen, W. *et al.* Improved analyses of GWAS summary statistics by reducing data heterogeneity and errors. *Nat. Commun.* **12**, 7117 (2021).
84. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
85. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
86. Yengo, L. *et al.* A saturated map of common genetic variants associated with human height. *Nature* **610**, 704–712 (2022).
87. Wright, S. The genetical structure of populations. *Ann. Eugen.* **15**, 323–354 (1951).
88. Hudson, R. R., Slatkin, M. & Maddison, W. P. Estimation of levels of gene flow from DNA sequence data. *Genetics* **132**, 583–589 (1992).
89. Nock, M. K. *et al.* Cross-national prevalence and risk factors for suicidal ideation, plans and attempts. *Br. J. Psychiatry* **192**, 98–105 (2008).
90. Coleman, J. R. I. *et al.* The genetics of the mood disorder spectrum: Genome-wide association analyses of more than 185,000 cases and 439,000 controls. *Biol. Psychiatry* **88**, 169–184 (2020).
91. R Core Team. R: A Language and Environment for Statistical Computing. https://www.R-project.org (2025).
92. Lee, S. H., Goddard, M. E., Wray, N. R. & Visscher, P. M. A better coefficient of determination for genetic profile analysis: A better coefficient of determination. *Genet. Epidemiol.* **36**, 214–224 (2012).
93. Angelo Canty & B. D. Ripley. boot: Bootstrap R (S-Plus) Functions. Preprint at (2024).
94. A. C. Davison & D. V. Hinkley. Bootstrap Methods and Their Applications. Preprint at (1997).
95. Robin, X. *et al.* pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* vol. 12 77 Preprint at (2011).

96. Koromina, M. *et al.* Fine-mapping genomic loci refines bipolar disorder risk genes. *Nat. Neurosci.* **28**, 1393–1403 (2025).

97. Gazal, S. *et al.* Functional architecture of low-frequency variants highlights strength of negative selection across coding and non-coding annotations. *Nat. Genet.* **50**, 1600–1607 (2018).

98. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15545–15550 (2005).

99. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **25**, 25–29 (2000).

100. Koopmans, F. *et al.* SynGO: An evidence-based, expert-curated knowledge base for the synapse. *Neuron* **103**, 217–234.e4 (2019).

101. Hatoum, A. S. *et al.* Psychiatric genome-wide association study enrichment shows promise for future psychopharmaceutical discoveries. *Commun. Med. (Lond.)* **5**, 176 (2025).